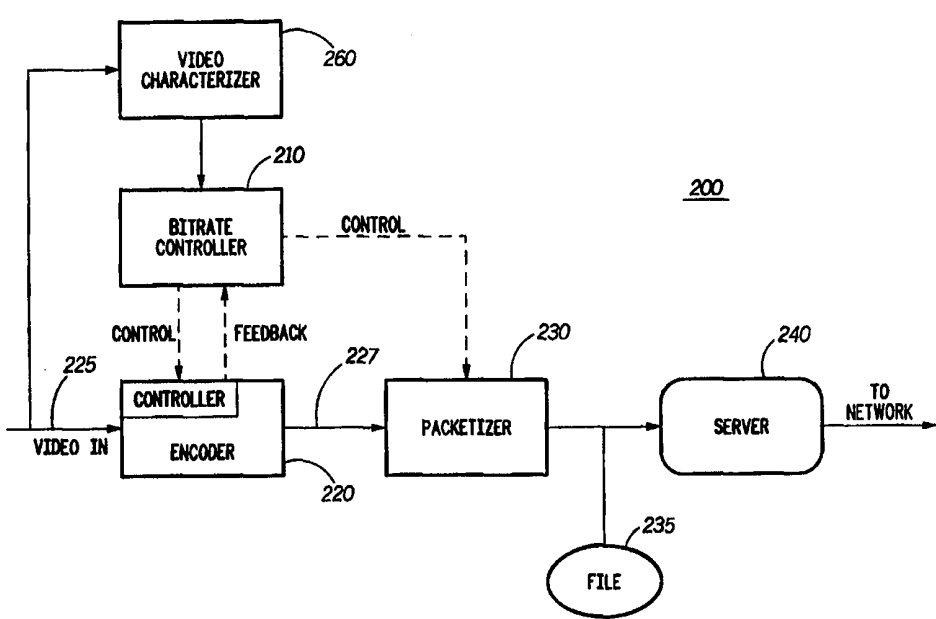




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : H04N 7/12</p>	<p>A1</p>	<p>(11) International Publication Number: <b>WO 99/43157</b> (43) International Publication Date: 26 August 1999 (26.08.99)</p>
<p>(21) International Application Number: PCT/US99/02411 (22) International Filing Date: 4 February 1999 (04.02.99) (30) Priority Data: 09/026,280 19 February 1998 (19.02.98) US (71) Applicant: MOTOROLA INC. [US/US]; 1303 East Algonquin Road, Schaumburg, IL 60196 (US). (72) Inventor: WANG, Feng, Chi; 12 Manhattan Road, Newton, MA 02165 (US). (74) Agents: WOOD, J., Ray et al.; Motorola Inc., Intellectual Property Dept., 1303 East Algonquin Road, Schaumburg, IL 60196 (US).</p>		<p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p><b>Published</b> <i>With international search report.</i></p>

(54) Title: SYSTEM AND METHOD FOR NON-CAUSAL ENCODING OF VIDEO INFORMATION FOR IMPROVED STREAMING THEREOF



(57) Abstract

A system (200) and method of non-causal encoding of video information (225), the video information (225) including a plurality of video segments, which involves: analyzing (260) the video information (225) to obtain global characterization information about the video information (225) and segment characterization information (265) about the plurality of video segments; and encoding (220) the video segments based on the global characterization information and the segment characterization (265) of the video segment being encoded (227).

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<b>AL</b>	Albania	<b>ES</b>	Spain	<b>LS</b>	Lesotho	<b>SI</b>	Slovenia
<b>AM</b>	Armenia	<b>FI</b>	Finland	<b>LT</b>	Lithuania	<b>SK</b>	Slovakia
<b>AT</b>	Austria	<b>FR</b>	France	<b>LU</b>	Luxembourg	<b>SN</b>	Senegal
<b>AU</b>	Australia	<b>GA</b>	Gabon	<b>LV</b>	Latvia	<b>SZ</b>	Swaziland
<b>AZ</b>	Azerbaijan	<b>GB</b>	United Kingdom	<b>MC</b>	Monaco	<b>TD</b>	Chad
<b>BA</b>	Bosnia and Herzegovina	<b>GE</b>	Georgia	<b>MD</b>	Republic of Moldova	<b>TG</b>	Togo
<b>BB</b>	Barbados	<b>GH</b>	Ghana	<b>MG</b>	Madagascar	<b>TJ</b>	Tajikistan
<b>BE</b>	Belgium	<b>GN</b>	Guinea	<b>MK</b>	The former Yugoslav Republic of Macedonia	<b>TM</b>	Turkmenistan
<b>BF</b>	Burkina Faso	<b>GR</b>	Greece			<b>TR</b>	Turkey
<b>BG</b>	Bulgaria	<b>HU</b>	Hungary	<b>ML</b>	Mali	<b>TT</b>	Trinidad and Tobago
<b>BJ</b>	Benin	<b>IE</b>	Ireland	<b>MN</b>	Mongolia	<b>UA</b>	Ukraine
<b>BR</b>	Brazil	<b>IL</b>	Israel	<b>MR</b>	Mauritania	<b>UG</b>	Uganda
<b>BY</b>	Belarus	<b>IS</b>	Iceland	<b>MW</b>	Malawi	<b>US</b>	United States of America
<b>CA</b>	Canada	<b>IT</b>	Italy	<b>MX</b>	Mexico	<b>UZ</b>	Uzbekistan
<b>CF</b>	Central African Republic	<b>JP</b>	Japan	<b>NE</b>	Niger	<b>VN</b>	Viet Nam
<b>CG</b>	Congo	<b>KE</b>	Kenya	<b>NL</b>	Netherlands	<b>YU</b>	Yugoslavia
<b>CH</b>	Switzerland	<b>KG</b>	Kyrgyzstan	<b>NO</b>	Norway	<b>ZW</b>	Zimbabwe
<b>CI</b>	Côte d'Ivoire	<b>KP</b>	Democratic People's Republic of Korea	<b>NZ</b>	New Zealand		
<b>CM</b>	Cameroon	<b>KR</b>	Republic of Korea	<b>PL</b>	Poland		
<b>CN</b>	China	<b>KZ</b>	Kazakstan	<b>PT</b>	Portugal		
<b>CU</b>	Cuba	<b>LC</b>	Saint Lucia	<b>RO</b>	Romania		
<b>CZ</b>	Czech Republic	<b>LI</b>	Liechtenstein	<b>RU</b>	Russian Federation		
<b>DE</b>	Germany	<b>LK</b>	Sri Lanka	<b>SD</b>	Sudan		
<b>DK</b>	Denmark	<b>LR</b>	Liberia	<b>SE</b>	Sweden		
<b>EE</b>	Estonia			<b>SG</b>	Singapore		

## **System and Method for Non-Causal Encoding of Video Information for Improved Streaming Thereof**

### **Cross-Reference to Related Applications**

5           This application is related to the following U.S. Patent Applications, both of which are assigned to the assignee of this application and incorporated by reference herein:

          System and Device for, and Method of, Encoding Video Information for Improved Streaming Thereof

10           having Feng Chi Wang as an inventor, U.S. Pat. Apl. Ser. No. 08/885,076, filed June 30, 1997; and

          Improved Video Encoding System and Method having Feng Chi Wang and Manickam Sridhar as inventors, U.S. Pat. Apl. Ser. No. 08/711,702, filed September 6, 1996.

15

### **Field of invention**

          The invention generally relates to multimedia applications and, more particularly, to the encoding and streaming of video information over a communication network.

20

### **Background of Invention**

          Generally speaking, there are two modern approaches to "playing-back" multimedia information located at a remote location, such as playing-back a "video clip" on the Internet. The first approach is to have a client node download a file having the video information from a corresponding "website," or server node, and to then play-back the information, once the file has been completely transferred. The second approach is to have the server node "stream" the information to the client node so that the client may begin play-back soon after the information starts to arrive. Because the streaming approach does not suffer from the long start-up delays inherent in the downloading approach, it is believed to be preferable in certain regards.

30

          It is believed that a substantial number of remote access users, such as Internet users, access the network via a voiceband modem. To this end, various communication standards have been proposed. H.261 and H.263, for example, each specify a coded representation that can be used for compressing video at low bitrates. (See ITU-T Recommendation H.263 of 2 May 1996, which is hereby incorporated by reference in its entirety.)

35

Because typical voiceband modems have maximum data rates of less than 56 Kb/s, the quality of a streamed play-back depends on how effectively the channel is used. TrueStream Streaming Software, version 1.1, for example, keeps the channel at full utilization to improve the play-back's appearance.

5 (TrueStream Streaming Software, version 1.1, is available from Motorola, Inc.)

In short, with version 1.1 of the TrueStream Streaming Software, a target data rate is first selected, for example, 20 Kb/sec for a 28.8 Kb/sec modem. (The other 8.8 Kb/sec of bandwidth is saved for audio information and packet overhead.) If a sequence of video frames is to be encoded and because of its  
10 inherent informational content the streaming of the encoded data would require a data rate higher than the target rate to maintain a certain image quality level, then the TrueStream system adjusts certain encoding parameters to reduce image quality and compress the image so that encoded frames fit into the channel bandwidth. On the other hand, if a sequence of video frames is to be encoded  
15 such that the streaming of it would not fully utilize the channel, the TrueStream system applies a "use it or lose it" approach and adjusts certain encoding parameters to improve the video quality and so that the channel capacity is used. For some images, e.g. a still blank screen, the expenditure of additional bits will have little or no improvement on the image quality. In these cases, the "use it or  
20 lose it" approach wastes bits. The consequence of the above is that in the former case the sequence will be played-back with pictures having a relatively coarser level of detail and a relatively smaller frame rate, and that in the latter case, the sequence will be played-back with pictures having a finer level of detail and a higher frame rate.

25

### **Brief Description of the Drawings**

In the Drawing,

Figure 1 shows a standard source encoder 100 having a known architecture;

30 Figure 2 shows a system architecture of an exemplary embodiment of the invention;

Figure 3 is a flowchart of characterization logic of an exemplary embodiment of the invention;

35 Figures 4A-G are a flowchart of bitrate controller logic of an exemplary embodiment of the invention.

### Detailed Description

The invention involves a system and method for non-causal encoding of video information for improved streaming thereof. Exemplary embodiments of the invention encode video information at variable bitrates and by doing so can utilize the channel in new and useful ways. For example, if the informational content of certain video frames can be encoded at a relatively low bitrate, the otherwise-unused channel capacity can be used to transmit information that will be needed for future video frames. This might be helpful in that a future portion of a video stream might, because of its informational content, require so much data that it otherwise is streamable only by sacrificing the quality of its play-back.

Exemplary embodiments first analyze and characterize the information to be encoded. The information is then encoded based on characterization information developed from analyzing the video. Thus, the encoding of a given frame depends not only on the characterization of the informational needs of the data corresponding to the given frame but to past and future frames, relative to the given frame. By characterizing the information before encoding, the exemplary embodiment gains knowledge about the clip and can make better decisions in allocating bandwidth resources. The characterizing of future information and use of the characterization of future video in making decisions for encoding of present video makes the system properly classified as "non-causal."

The exemplary embodiments are particularly concerned with video information encoded according to H.263. Thus, material aspects of H.263 are outlined below, followed by a description of the exemplary embodiments.

25

#### I. Outline of H.263

A video sequence is encoded and decoded as a sequence of frames or pictures. Each picture is organized as groups of blocks (GOBs), macroblocks, and blocks, and each picture may be of a variety of picture formats and subformats. In addition, each picture may be of the INTRA type, also known as an "I" frame, or the INTER type, which includes entities known as "P" frames and "PB" frames.

An "I" frame is independent in that it represents a complete image. Its encoding and decoding have no dependencies on prior frames. With "P" and "PB" frames, on the other hand, the encoding and decoding depends on prior and/or future frames. P and PB frames may be thought of as an encoded representation of the difference between one picture and other picture(s).

35

Figure 1 shows a standard source encoder 100. Coding controller 110, among other things, controls switches 120 and 130 and quantizer Q. In short, the controller 110 controls the sampling rate, or frame rate, of the video sequence by selecting particular frames of Video In, and it controls the level of detail of each encoded picture by supplying quantization parameters  $q_z$  to quantizer Q. The output information 'q' is a compressed version of a picture: an I frame is a compressed version of a current picture, or portion thereof, and a P or PB frame is a compressed version of difference information, representing the difference between the current picture, or portion thereof, and the last, or next in the case of B-frames.

For I frame blocks, coding controller 110 connects switch 120 to input 121 and switch 130 to input 131. Thus, Video In is connected to the transform block T, which processes the data according to a known discrete cosine transform (DCT). The transformed data, known as transform coefficients, are then quantized by quantizer Q, and the quantized information 'q' is received by inverse quantizer  $Q^{-1}$ . The inverse quantized information, in turn, is received by inverse transform  $T^{-1}$ , and the inverse transformed information is received by summation node 140. Consequently, summation node 140 adds unconnected input 131 and the reconstituted picture information from the inverse transform  $T^{-1}$ , the reconstituted picture information being the picture Video In after it has been transformed, quantized, inverse quantized, and inverse transformed. The output of the summation node 140 is thus the reconstituted picture information, which is stored in picture memory P.

For P or PB type blocks, coding controller 110 controls switches 120 and 130 to connect to inputs 122 and 132, respectively. Difference node 150 produces the difference between the current block, Video In, and the prior, reconstituted block, provided by picture memory P. This difference information is then transformed (T), quantized (Q), inverse quantized ( $Q^{-1}$ ), and inverse transformed ( $T^{-1}$ ) analogously to that described above. The information provided by inverse transform ( $T^{-1}$ ), in this arrangement, however, is not the reconstituted picture, but rather the reconstituted difference information originally provided by difference node 150. Summation node 140 adds the reconstituted difference information to the prior block, provided by input 132, to yield a reconstituted version of the current picture, which is stored in picture memory P.

The compression gains achieved by the above arrangement result from the statistical nature of video information and from the quantization rules. In particular, a given pixel in one frame is likely to be the same or nearly the same as

a corresponding pixel of a prior frame. Moreover, pixels having no difference from one picture to the next tend to run contiguously so that many adjacent pixels might be identical to the corresponding pixels of a prior frame. The H.263 encoding methods address the above with a variety of techniques that can be summarized by stating that the more probable video information events require less encoded bits than the less probable events, to maintain a certain image quality. (See ITU-T Recommendation H.263 of 2 May 1996, at 25-27, variable and fixed length codes for transform coefficients.)

The amount of bits needed to represent a picture or block depends primarily on three things: (a) whether the blocks are being encoded as INTER or INTRA type; (b) the informational nature of a block in relation to a prior block; and (c) the level of quantization used in the encoding. Typically, INTER-based encoding requires less bits than INTRA-based encoding for a given frame, and coarser detail requires less encoded bits than finer detail, everything else being equal. Thus, the selection of INTER or INTRA and the level of quantization are the primary independent variables affecting the amount of data required to encode a given picture, Video In.

## II. Overview of a System and Method for Non-Causal Encoding of Video Information for Improved Streaming Thereof

Figure 2 shows an exemplary system 200 for encoding and packetizing files of variable bitrate video data. Video In, i.e., unencoded data, is received by encoder 220 and video characterizer 260. The characterizer 260 analyzes and characterizes the video portion. After, the video has been characterized, the encoder 220 encodes the information, under the control of coding control 225 and bitrate controller 210. The bitrate controller 210, in turn, acts in response to characterization information 265 provided by characterizer 260. The encoded data 227 is received by packetizer 230, which is partially controlled by bitrate controller 210, and which packetizes the information into a specified format to create a file 235. File 235 may then be used by server logic 240 to stream the information over the network 250.

The above embodiment thus forms a two-pass approach to encoding. The first pass analyzes and characterizes the video information to be encoded. The second pass encodes the video information to be encoded, adjusting encoding parameters based on the characterization of the video.

a. Video Characterizer

An exemplary characterizer 260 analyzes and characterizes the entire set of data to be encoded before any encoding operations begin. The characterization provides information indicative of the average motion and the average error of the video, as well as indications of scene changes. Both "motion" and "error" are characteristics known in the video encoding art. A "scene change" is inferred whenever the error of one frame relative to a prior frame is so large that the probability that the two frames represent a change in scenes is high.

Figure 3 is a flowchart of exemplary characterization logic. The logic starts in step 300 and proceeds to step 310, which initializes characterization parameters and divides the video to be encoded into N segments. Each segment corresponds to a predetermined time duration, e.g., 3 seconds of playback, which in turn depends on the reaction time of the codecs involved. Thus, in an encoding arrangement targeting 30 frames/second, a segment would correspond to 90 frames of video. As part of the dividing step, the last segment (N+1) has a high likelihood of corresponding to fractional portion of 3 seconds and is treated as such by default. It must be noted that this segment size is just exemplary, and various segment sizes may be used with as few as two frames of video per segment.

The logic proceeds to step 320 to determine whether the segment pointer is pointing to the last segment. If so, the logic proceeds to step 330 which characterizes the segment with default values (more below) and then proceeds to step 399 to end the logic flow.

If step 320 determines that the last segment is not being pointed to, the logic proceeds to step 340. In step 340, the segment is considered in sub-segments with each sub-segment corresponding to what will eventually be a frame of data. Thus, if the targeted encoding rate is 30 frames per second, a sub-segment will correspond to 1/30th of a second of video. Starting with the second such sub-segment and for each subsequent sub-segment of the segment, e.g., sub-segment 2 through 90, the logic calculates a Motion Parameter (MP) and an Error Parameter (EP) using largely conventional techniques such as arithmetic difference and city-block distance.

Namely, to determine MP and EP for a sub-segment, an exhaustive search is performed to find the Motion Vector that results in the smallest Y-Luma differential coding error with respect to a prior frame. This is done for each macroblock-corresponding portion of the sub-segment. The Motion Vectors are limited to [-16, 15.5] and are restricted to data that is completely within the video



frame boundaries, i.e., the Unrestricted Motion Vector Annex (of H.263) is not allowed. An exemplary way of determining the Y-Luma error is calculated by averaging over the absolute values of the differences of all pixels in the MacroBlock (i.e., this is not a RMS error). The resulting Motion Vector pair [dx,dy] is used to calculate MP by summing the absolute values of dx and dy (city block distance). EP is set to the corresponding Y-Luma error summation.

The logic then proceeds to step 350 which calculates an Average Motion Parameter (AMP) and an Average Error Parameter (AEP) for the segment by considering the MPs and EPs of the sub-segments.

The logic then proceeds to step 360 which determines whether any of the EPs for the sub-segments exceeds a threshold value that would correspond to a scene change. Scene changes are important from an informational standpoint and thus warrant the use of "extra bandwidth" to encode the video. Each EP must be considered to detect a scene change, i.e., a large error from one frame to the next, because there is no guarantee that a scene change could be detected from analyzing AEP alone.

The logic then proceeds to step 370 which updates a data structure used to hold information characterizing the video. Among other things, the structure holds AEP and AMP values for each segment and includes a marker to indicate whether a given segment has a scene change, i.e., a relatively high EP for at least one of the sub-segments in the segment.

The logic proceeds back to step 320 to determine whether more segments need to be characterized, as explained above.

#### b. Encoder and Controller

Encoder 220 includes the elements of Figure 1, except that the coding controller 110 of Figure 1 is divided into controller 225 and bitrate controller 210, discussed below. Controller 225 includes control features known in the art as well as control features described in the related applications identified and incorporated above. (The application entitled Improved Video Encoding System And Method includes features that would control the bitrate of the data stream in a mutually-exclusive manner to that covered by the bitrate controller 210. Under one embodiment of the invention, a user input allows selection between the two.

c. Bitrate Controller

Figures 4A-G are flowcharts of exemplary bitrate controller logic. Briefly, the logic operates in the following manner. A target bitrate and quantization level are received from a user. Each segment, e.g., 3 seconds of video, is then compared to some global characteristic information. Thus, it may be determined whether a given segment has more motion than average or more error than average. Depending on the comparison the bitrate and quantization parameters may be adjusted relative to the targets. Thus, a segment having a relatively large amount of motion may be allocated more bits of encoded information than a segment having less motion. After a segment is analyzed in comparison to the global information, and consequently after the encoding parameters are possibly adjusted in response thereto, the segment is encoded using conventional techniques. Thus, the encoding of a given segment depends on characteristics of prior and future segments to be encoded.

The logic starts at step 400 and proceeds to step 402 in which a target bitrate (TBitrate) and a target quantization level (TQuant) are received as user inputs. For example, on a targeted 28.8 Kb/s connection, TBitrate might be set at 20 Kb/s to allow adequate resources for audio data or the like. Under H.263, Quant may vary from 1-31 with lower numbers corresponding to finer detail and consequently requiring more bits to encode the data.

The logic proceeds to step 404 in which the first segment is encoded using TBitrate and TQuant as the encoding parameters. The encoding of video data, responsive to the above parameters, is conventional.

The logic proceeds to step 406 in which a Global Average Error Parameter (GAEP) and a Global Average Motion Parameter (GAMP) are calculated. This is done by averaging the AEPs and the AMPs, respectively, of the N-1 segments between the first segment and the fractional segment. GAEP and GAMPs thus provide characteristic information of the entire video clip to be encoded.

The logic proceeds to step 408 in which the segment pointer (SP) and the Bitcredit variable (Bitcredit) are initialized. SP is initialized to the second segment and Bitcredit is initialized to 0. As will be explained below, Bitcredit effectively keeps a running total of the otherwise unused bandwidth which may be used for encoding video, if needed.

The logic proceeds to step 410 in which it is determined whether the AEP of the current segment being encoded (i.e., AEP[SP]) has twice the global average error (i.e., energy) of the whole clip (GAEP). The step also determines whether the AEP of the current segment is 1.5 times the global average of the whole clip

and whether the current segment (i.e., SP) has a scene change marker set from the characterization phase.

If either of the above are true, then the logic proceeds to step 412 which sets the variable Bitrate equal to the TBitrate and the Bitcredit (which initially starts at zero, but which as explained below can grow as a result of frames not needing many bits for encoding thereof).

The logic proceeds to step 414 in which it is determined whether the AMP of the current segment has 1.5 times the global average motion of the whole clip (GAMP). If so, the logic proceeds to step 416 in which the variable Quant is set to the target Quant (TQuant) plus three; thus setting the quantization parameter to correspond to coarser detail to better encode the relatively higher motion segment. If step 414 determines that the current segment is not a relatively high motion segment, as compared to the global average, the logic proceeds to step 418, in which it is determined whether the current segment is a relatively low motion segment. In particular, the AMP of the current segment is compared to 0.5 times the GAMP. If the AMP is less, corresponding to a relatively low motion segment, the logic proceeds to step 420 in which the quantization variable Quant is adjusted to correspond to finer detail; in this instance Quant equals TQuant minus three.

The logic then proceeds to step 468 (FIG. 4G) in which the current segment's frames are encoded using conventional encoding logic that is responsive to the encoding variables Quant and Bitrate (e.g., version 2.0 of Telenor's public domain software, or the logic of the related, pending applications identified above).

The logic then proceeds to step 470 in which the segment pointer is incremented and then to step 472 in which it is determined whether the segment pointer is pointing to the last whole segment.

If the segment pointer is pointing to the last whole segment, the last segment of the clip, that is, the fractional segment, is encoded using the target quantization (TQuant) and the target Bitrate (TBitrate) in step 474 and the logic ends in step 499.

If the segment pointer is not pointing to the last segment, the logic loops back to step 410 described above.

If step 410 determines that the current segment does not have a relatively high error parameter (i.e., not high energy), the logic proceeds to step 422, FIG. 4C. (The exemplary rules for determining relatively high error parameter were discussed above, when first discussing step 410)

Step 422 determines whether the current segment has a relatively high error parameter, but not as high as that described above. In particular, step 422 compares the AEP of the current segment to 1.5 times the global AEP and also compares the AEP to the global AEP if the current segment also includes a scene change. (Again, a scene change may not necessarily require a high AEP, but nonetheless corresponds to at least some of the frames having high error parameters) If the comparison determines that the current frame has this second stage of relatively high error parameter, the logic proceeds to step 424.

In step 424, Bitrate is set to TBitrate plus one-half of the Bitcredit. Thus, only some of the accumulated Bitcredit is allocated for the segments having a relatively high error parameter but which is less high than the first tier discussed above.

The logic proceeds to step 426 in which the AMP is compared to 1.5 times the GAMP, and in which if the AMP is larger the logic proceeds to step 428 to adjust Quant to TQuant plus two.

The logic proceeds to step 430 in which the AMP is compared to 0.5 times the GAMP, and in which if the AMP is smaller the logic proceeds to step 432 to adjust Quant to TQuant minus two. The logic then proceeds to step 468, FIG. 4G, described above.

If step 422 determines that the AEP and scene change marker do not indicate that the current segment is a second tier, relatively high error parameter segment, then the logic proceeds to step 434, FIG. 4D.

Step 434 determines whether or not the current segment has a scene change marker. As mentioned above, a segment having a scene change need not have as high an AEP to qualify for additional bandwidth allocation.. If the current segment does have a scene change, Bitrate and Quant are set to their target values in step 436 and the logic proceeds to step 468, as described above. If the current segment does not contain a scene change the logic proceeds to 438, FIG. 4E.

Step 438 determines whether the current segment has a relatively low error parameter, in particular whether AEP is 0.5 times the global average error parameter (GAEP). If the AEP is less, the logic proceeds to step 440.

In step 440, Bitrate is set to 0.5 times TBitrate, the underlying principle being that a relatively lower energy segment should require a lower bitrate for the encoding thereof. Correspondingly, the use of a lower bitrate correlates to otherwise available bandwidth not being used for the current segment. Thus, Bitcredit is adjusted upward by adding 0.5 times TBitrate to Bitcredit.

The logic proceeds to step 442 in which the AMP of the current segment is compared to 1.5 times the GAMP. If the AMP is larger, indicating that the current segment has relatively high motion compared to the whole clip, the logic proceeds to step 444 which adjusts Quant to be 2 higher, and thus coarser, than the target quantization TQuant. If the AMP is lower, the logic proceeds to step 446.

In step 446, the AMP is compared to 0.5 times the GAMP. If the AMP is lower, indicating that the current segment has relatively low motion compared to the whole clip, the logic proceeds to step 448 which adjusts Quant to be 2 lower, and thus finer, than the target quantization TQuant. If the AMP is higher, the logic proceeds to step 450.

In step 450, the AMP is compared to 0.1 times the GAMP. If the AMP is lower, indicating that the current segment has the relative lowest motion compared to the whole clip, and can be assumed to be a still frame, the logic proceeds to step 448 which adjusts Quant to be 5 lower, and thus finer, than the target quantization TQuant. The logic then proceeds to step 468, described above.

If step 438 determines that the error parameter for the current segment is not lower than 0.5 times the GAEP, the logic proceeds to step 454, FIG. 4F.

In step 454, the AEP is compared to 0.75 times the GAEP. This step, in conjunction with step 438, determines whether the AEP lay between 0.5 and 0.75 times the GAEP, or whether the AEP is greater than 0.75 times the GAEP. If the AEP is greater than 0.75 times the GAEP, the logic proceeds to step 456 which sets the quantization and bitrate encoding parameters, Quant and Bitrate, to the target values, set by the user. The logic would then proceed to step 468, described above.

If the AEP is below 0.75 the GAEP (which again in conjunction with step 438 means that the AEP is below 0.75 times the GAEP but above 0.5 times the GAEP), then the logic proceeds to step 458.

In step 458, Bitrate is set to 0.75 times the target bitrate, reflecting that this segment has relatively low error, or energy, compared to the average energy of the clip. This conservation of bitrate is also reflected in step 458 adjusting the Bitcredit upward by 0.25 times the bitrate. The logic then proceeds to step 460.

In step 460, the AMP is compared to 1.5 times the GAMP. If the AMP is greater, indicating that the current segment has relatively high motion compared to the whole clip, the logic proceeds to step 462 which adjusts Quant to be 2 higher, and thus coarser, than the target quantization TQuant. If the AMP is lower, the logic proceeds to step 464.

In step 464, the AMP is compared to 0.5 times the GAMP. If the AMP is lower, indicating that the current segment has relatively low motion compared to the whole clip, the logic proceeds to step 466 which adjusts Quant to be 1 lower, and thus finer, than the target quantization TQuant. The logic then proceeds to step 468, described above.

It should be noted that this invention may be embodied in software and/or firmware stored on a computer useable medium, such as a computer disk or memory chip. The invention may also take the form of a computer data signal embodied in a carrier wave, such as when the invention is embodied in software/firmware which is electrically transmitted, for example, over the Internet.

The present invention may be embodied in other specific forms without departing from the spirit or essential characteristics. The described embodiments are to be considered in all respects only as illustrative and not restrictive. The scope of the invention is, therefore, indicated by the appended claims rather than by the foregoing description. All changes which come within the meaning and range within the equivalency of the claims are to be embraced within their scope.

20

25

30

35

What is claimed is:

**Claims**

- 5           1. A method of non-causal encoding of video information, the video information including a plurality of video segments, the method comprising:
- analyzing the video information to obtain global characterization information about the video information and segment characterization information about the plurality of video segments; and
- 10           encoding the video segments based on the global characterization information and the segment characterization information of the video segment being encoded.
- 15           2. The method of claim 1 further including dividing the video information into the plurality (N+1) of video segments including a first video segment, a last video segment (N+1) and remaining (N-1) segments between the first and last video segments.
- 20           3. The method of claim 2 wherein the step of analyzing includes calculating an average motion parameter (AMP) and an average error parameter (AEP) for each of the video segments except the last video segment, determining if a scene change has occurred for each of the plurality of video segments and storing the AMP, AEP and scene change information as the segment characterization information for the plurality
- 25           of video segments.
- 30           4. The method of claim 3 wherein the step of analyzing further includes determining the global characterization information about the video information by calculating a global average error parameter (GAEP) and a global average motion parameter (GAMP) over the remaining (N-1) video segments.
- 35           5. The method of claim 4 wherein the step of encoding includes determining encoding parameters for the video segment of the remaining video segments being encoded based on the global characterization

information and the segment characterization information of the video segment being encoded.

5 6. The method of claim 5 wherein the step of determining includes determining a first encoding parameter, a bitrate parameter, by comparing the AEP of the segment to be encoded with the GAEP and by considering the scene change information associated with the video segment to be encoded.

10 7. The method of claim 6 wherein the step of determining further includes determining a second encoding parameter, a quantization parameter, by comparing the AMP of the segment to be encoded with the GAMP.

15 8. The method of claim 7 wherein the step of determining the bitrate parameter includes accumulating unused bitrate when the bitrate parameter for a video segment being encoded is set below a target bitrate due to low AEP for that video segment relative to the GAEP and utilizing the unused bitrate for video segments having an AEP exceeding the GAEP by a first predetermined amount or by a second, lesser  
20 predetermined amount coupled with a scene change by setting the bitrate parameter at a value which exceeds the target bitrate.

25 9. A system for non-causal encoding of video information, the video information including a plurality of video segments, the system comprising:

30 a video characterizer for analyzing the video information to obtain global characterization information about the video information and segment characterization information about the plurality of video segments;

a controller, responsive to the video characterizer, for determining encoding parameters based on the global and segment characterization information; and

35 an encoder, responsive to the controller, for encoding the video segments using the determined encoding parameters.

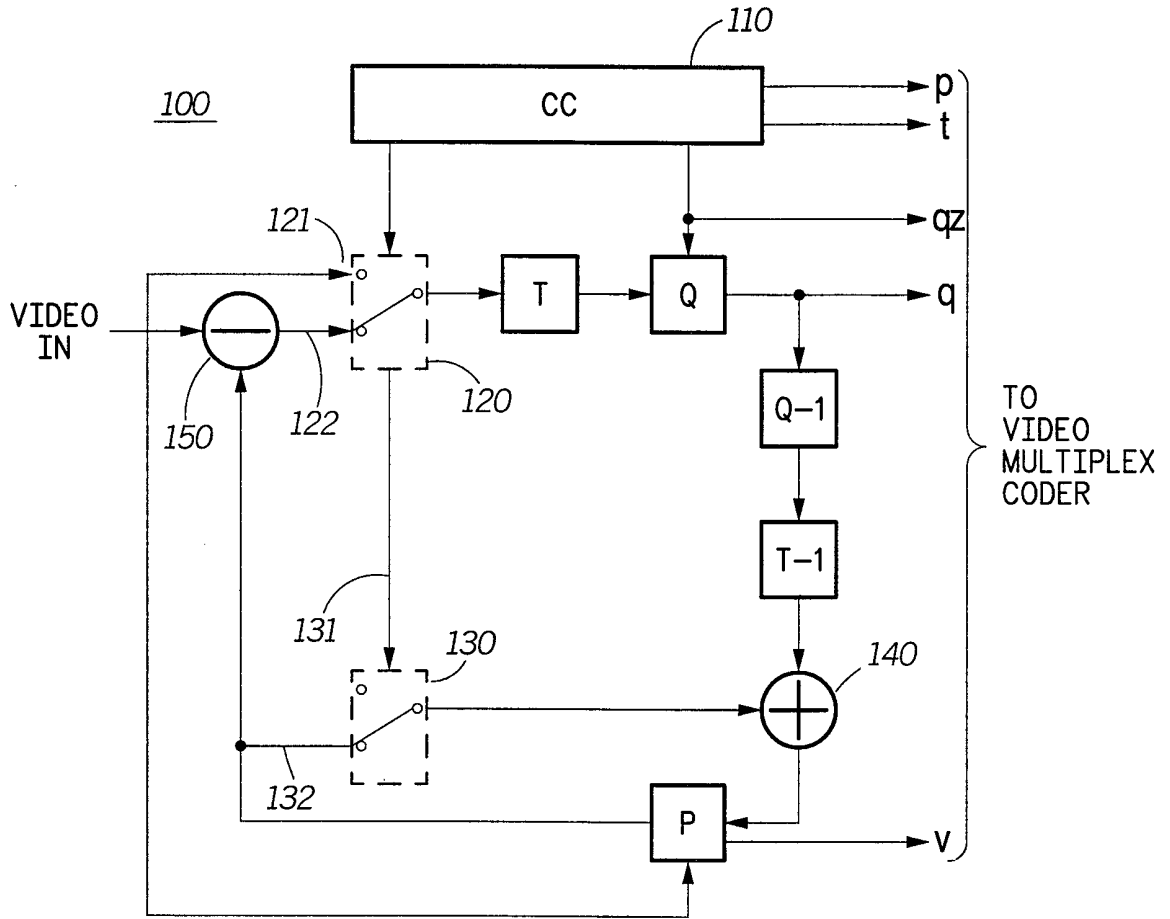


10. A computer useable medium having computer readable program code means embodied therein for non-causal encoding of video information, the video information including a plurality of video segments, comprising:

5 computer readable program code means for analyzing the video information to obtain global characterization information about the video information and segment characterization information about the plurality of video segments;

10 computer readable program code means, responsive to the computer readable program code means for analyzing, for determining encoding parameters based on the global and segment characterization information; and

15 computer readable program code means, responsive to the computer readable program code means for determining, for encoding the video segments using the determined encoding parameters.



TO VIDEO MULTIPLEX CODER

**FIG. 1**

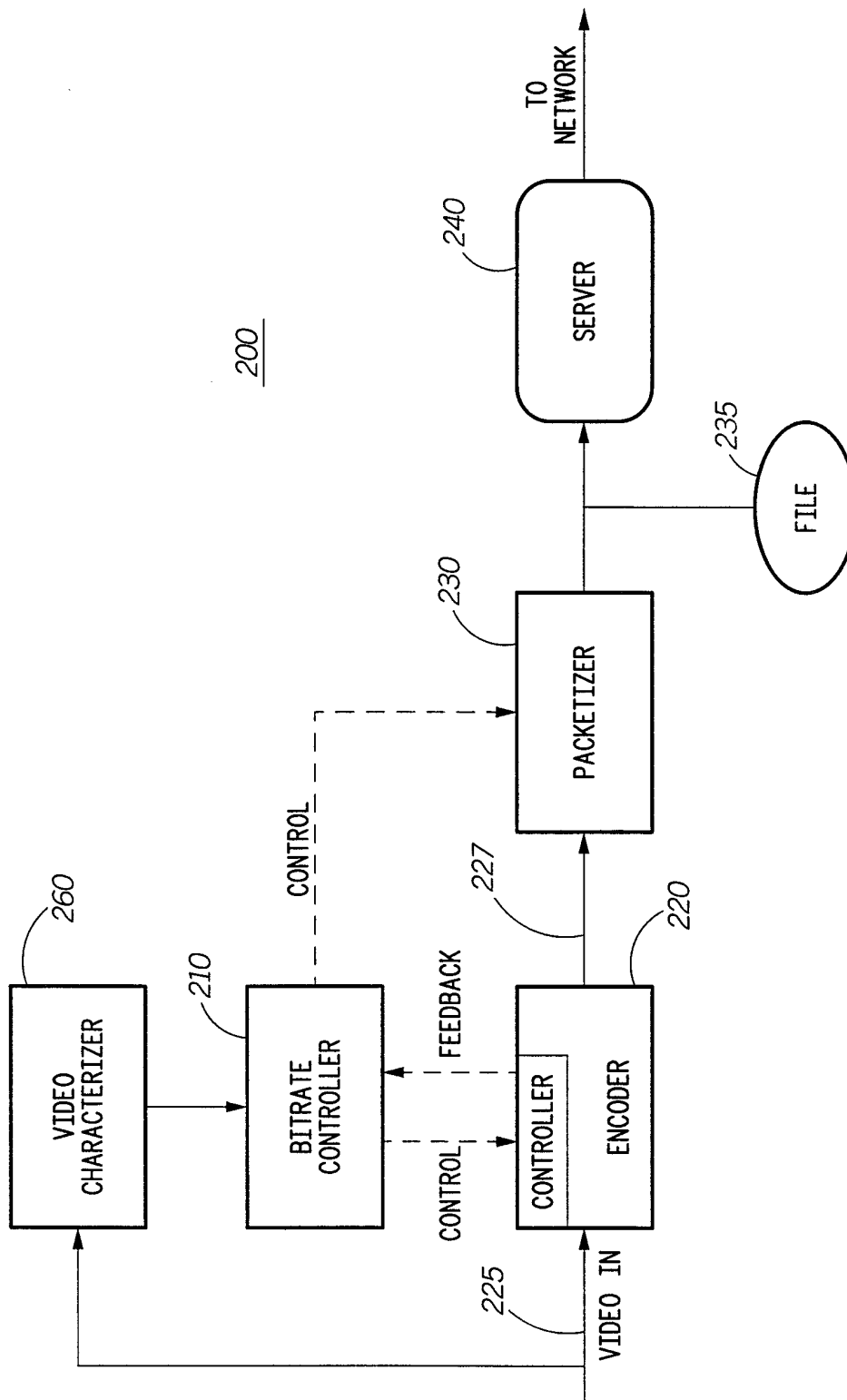
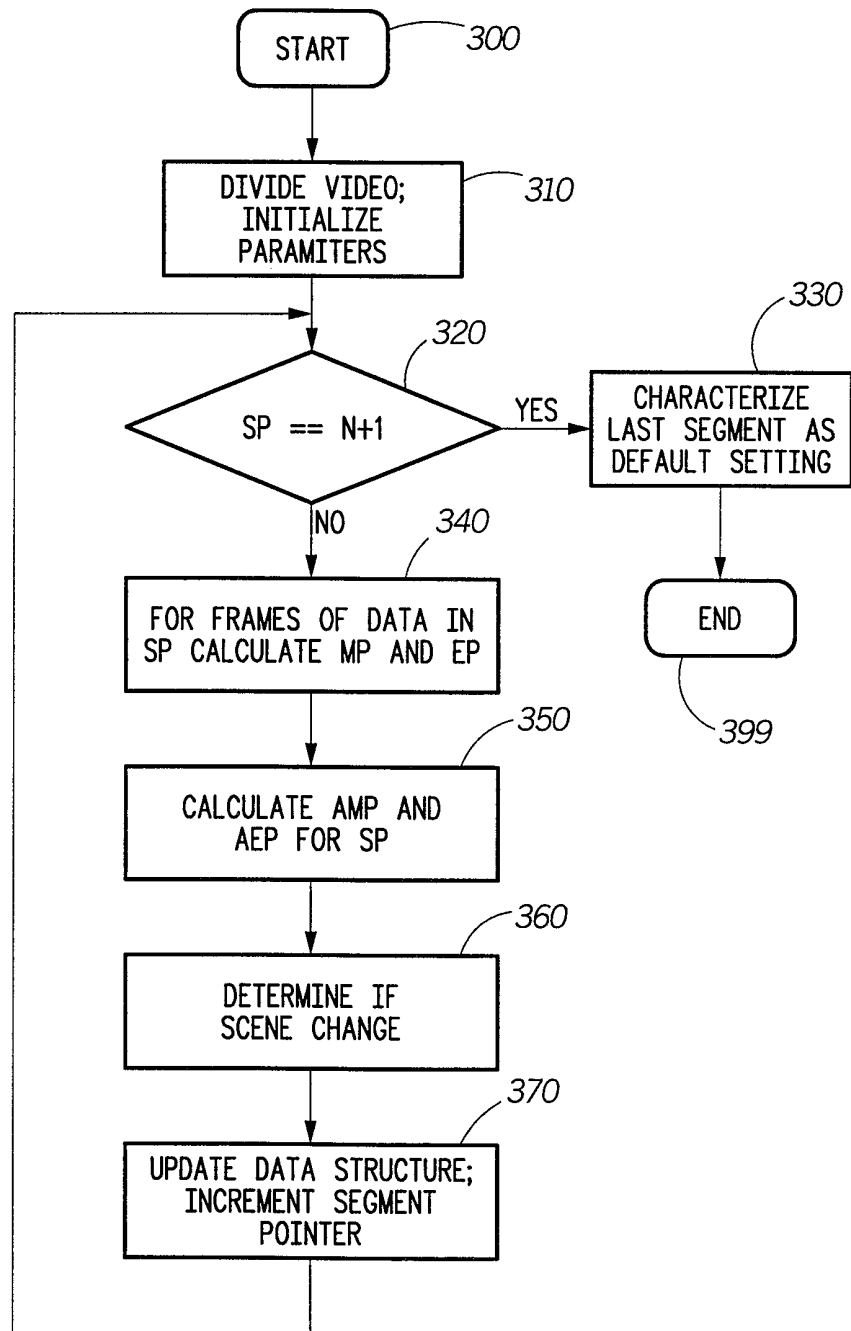
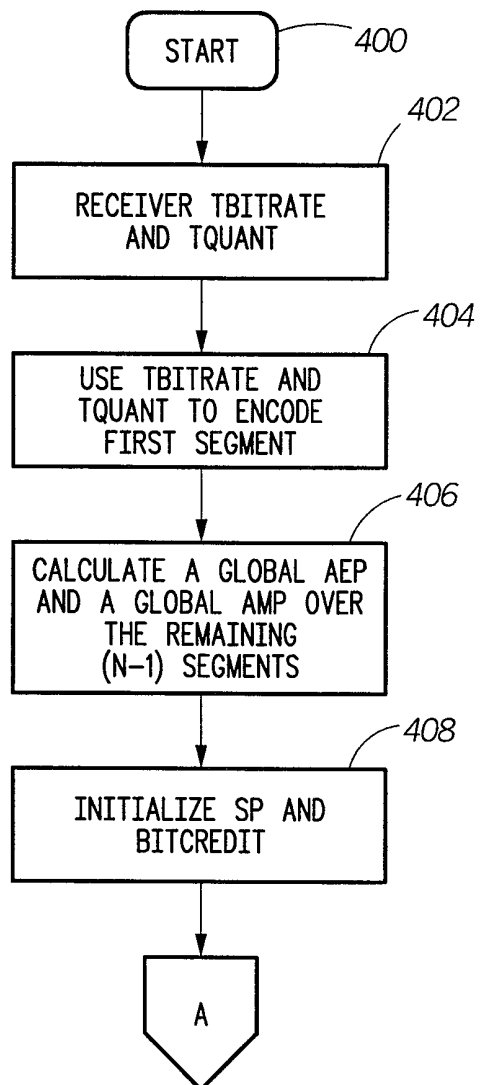


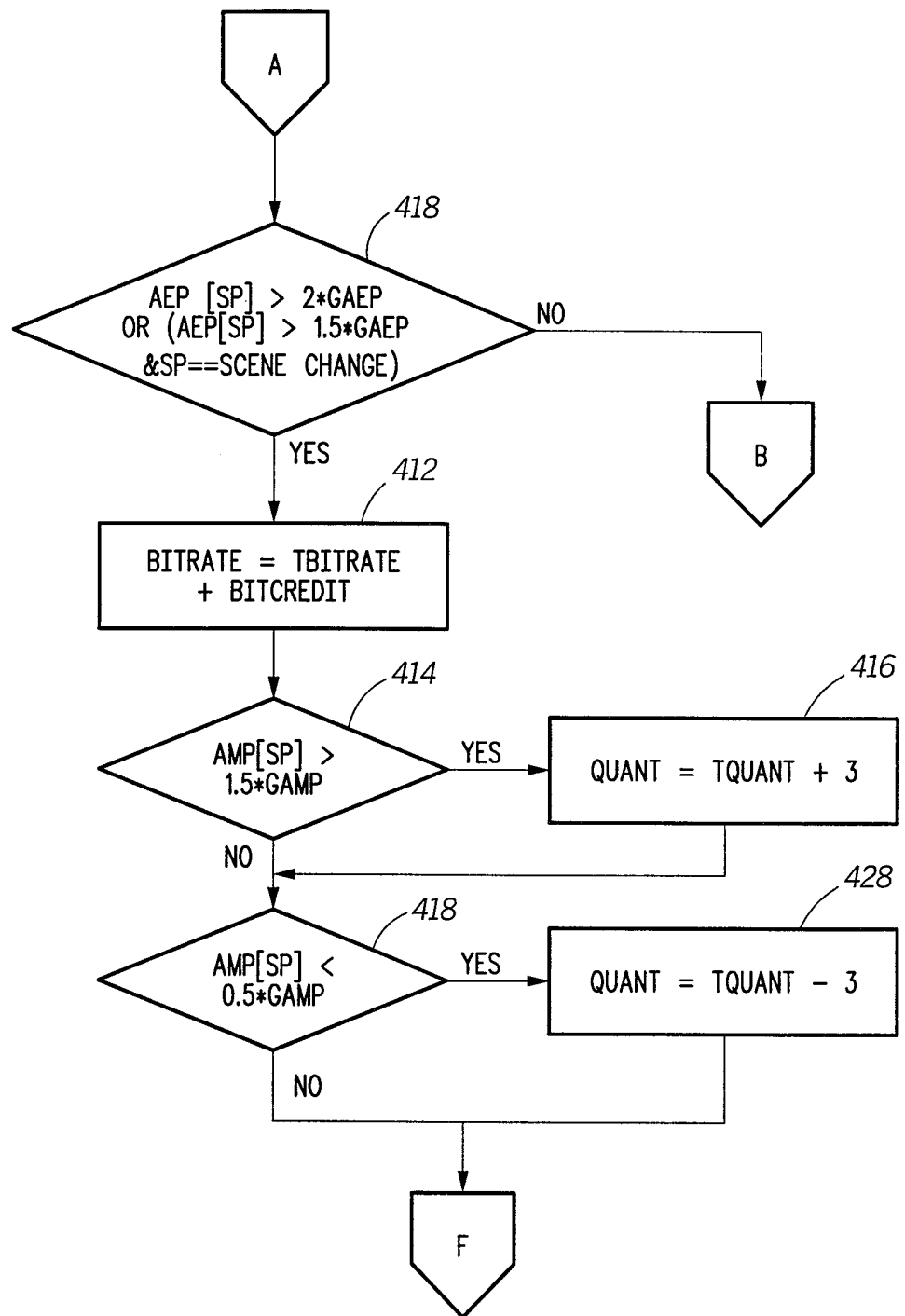
FIG. 2



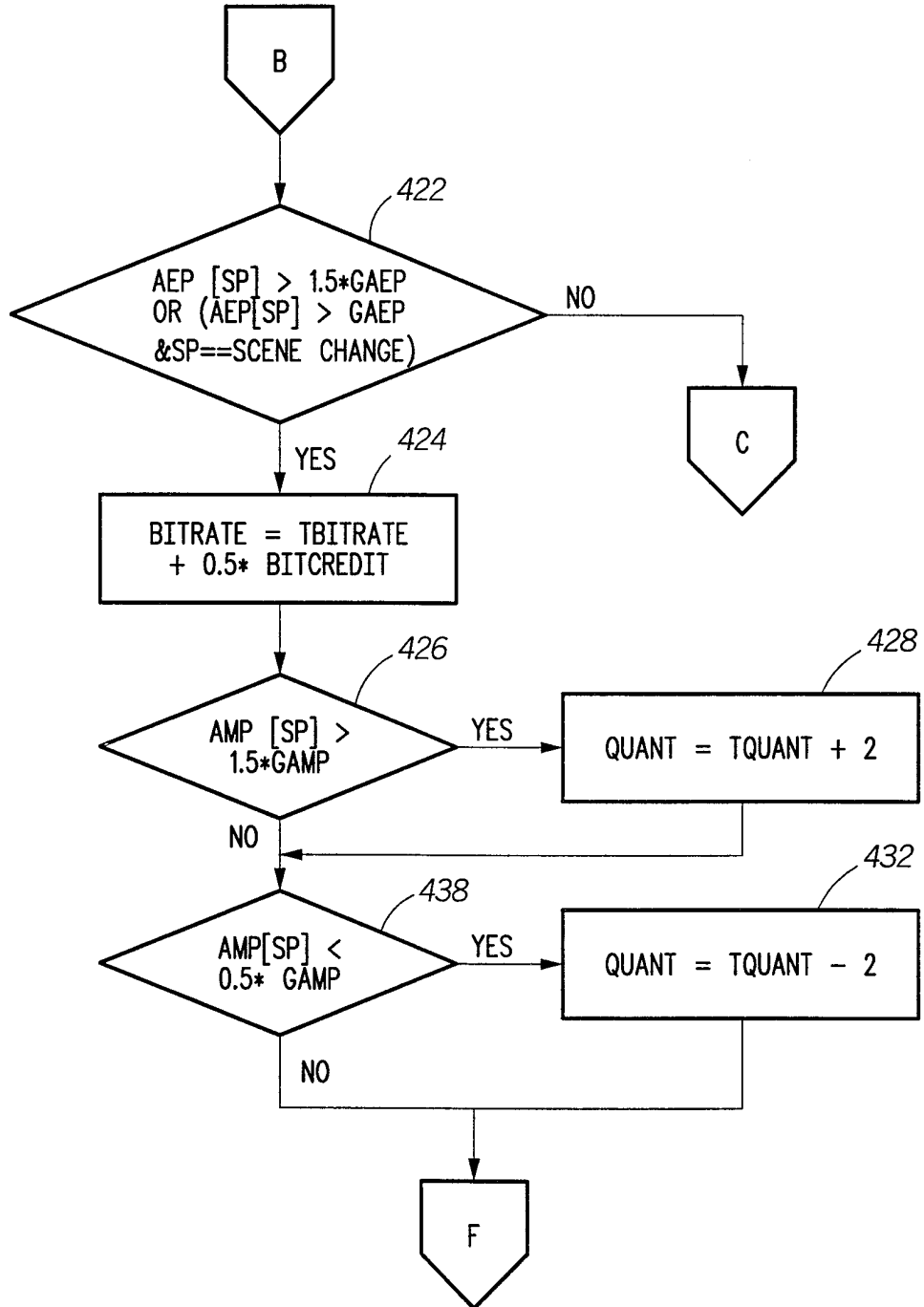
**FIG. 3**

4/10

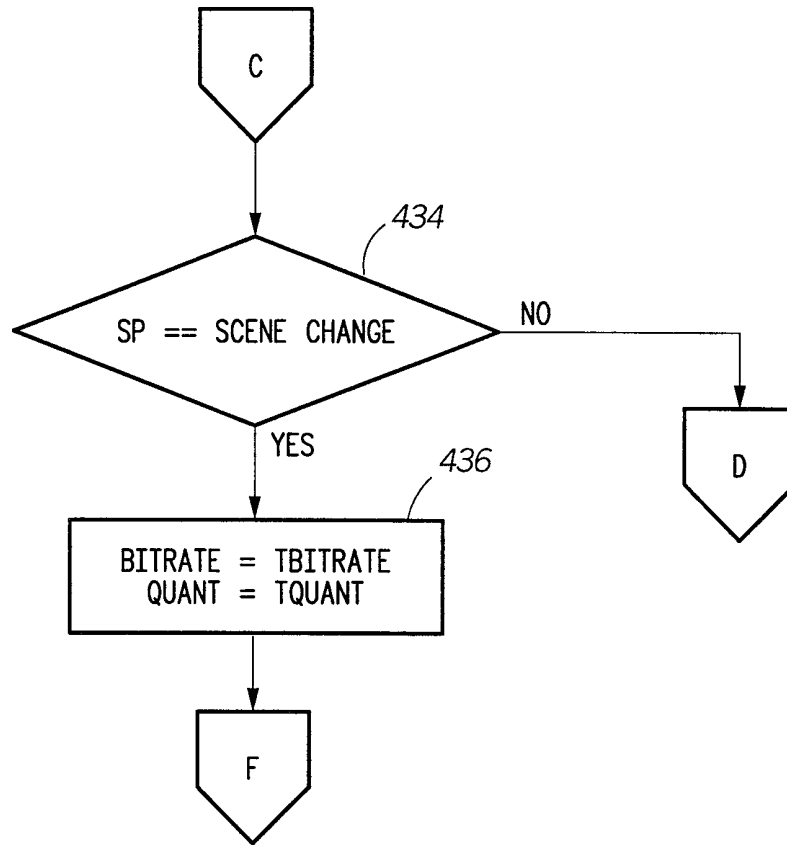
**FIG. 4A**



**FIG. 4B**

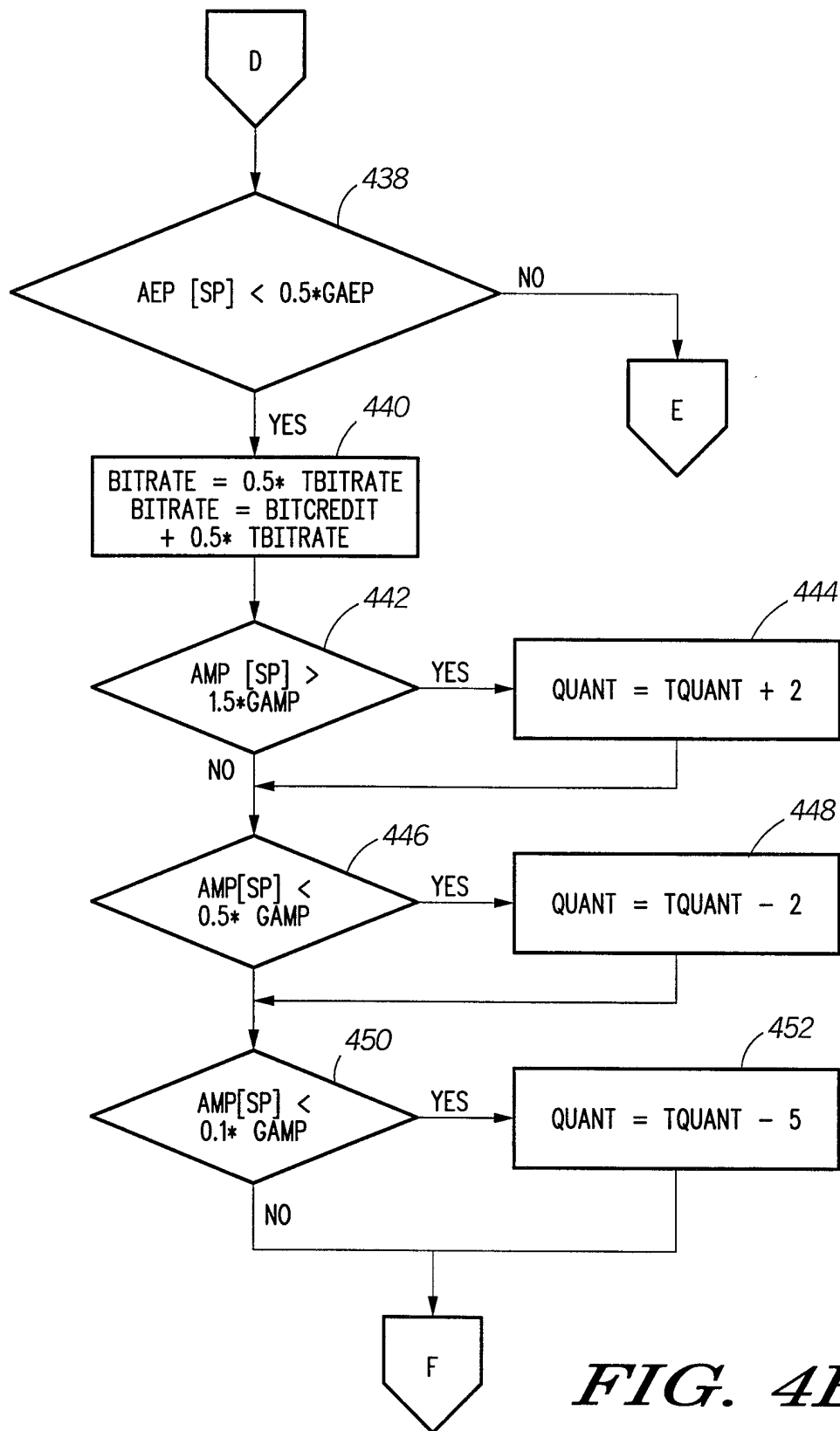


*FIG. 4C*

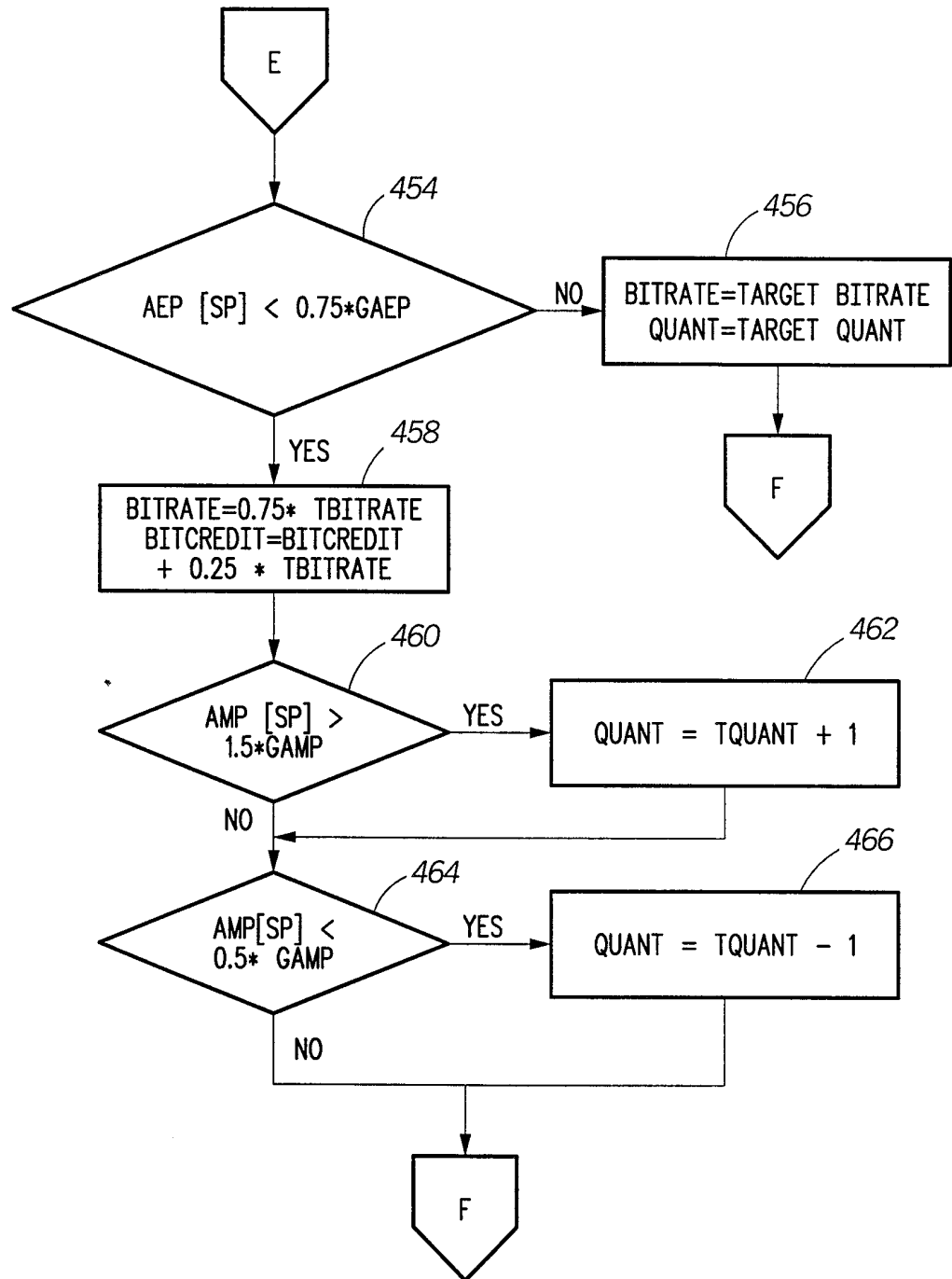


*FIG. 4D*

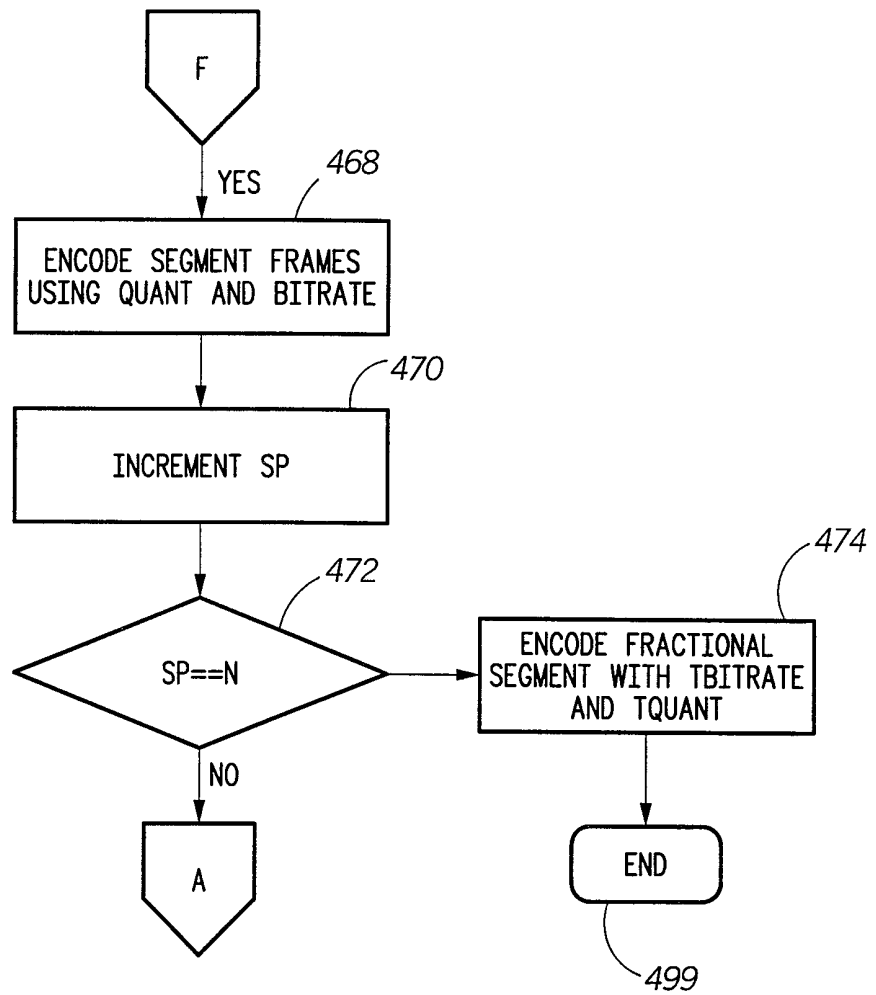




*FIG. 4E*



**FIG. 4F**



*FIG. 4G*

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/02411

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : H04N 7/12

US CL : 348/410, 416; 382/236

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 348/410, 416; 382/236

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P ----- Y, P	US 5,832,125 A (REESE et al.) 03 November 1998, col. 3, lines 4-10.	1, 9, 10 ----- 2-8
Y, P	US 5,774,593 A (ZICK et al.) 30 June 1998, col. 10, lines 48-67.	2-8
A	US 5,225,904 A (GOLIN et al.) 06 July 1993, col. 1, lines 18-46.	1-10
A	US 5,708,473 A (MEAD) 13 January 1998, col. 2, lines 39-67.	1-10
A	US 5,227,878 A (PURI et al.) 13 July 1993, col. 2, lines 48-68, col. 3, lines 1-6.	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

\*

Special categories of cited documents:

\*T\*

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*A\*

document defining the general state of the art which is not considered to be of particular relevance

\*X\*

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*E\*

earlier document published on or after the international filing date

\*L\*

document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*Y\*

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

\*O\*

document referring to an oral disclosure, use, exhibition or other means

\*P\*

document published prior to the international filing date but later than the priority date claimed

\*&amp;\*

document member of the same patent family

Date of the actual completion of the international search

05 APRIL 1999

Date of mailing of the international search report

03 JUN 1999

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

Tommy P. Chin

Joni Hill

Telephone No. (703) 305-4700